

The Power of Prediction: Beating Vegas with Simplicity

An Honors Thesis (HONR 499)

by

Derek Smith

Thesis Advisor

Prof. Gary Dean

December 2018

Expected Date of Graduation

December 2018

Abstract

The sport of modern football has been around for over one hundred years as entertainment for the American public. While the game plans have wildly changed, the spectators have not. Back in its inception, fans were betting on who they thought would be able to pull out the victory. As the game developed, so too did the betting markets. Now most bets on the modern day NFL are run through sports books in Las Vegas. The growth of statistics and computing power has made it difficult for the majority of people to consistently bet on NFL games and turn a profit, as Vegas has taken on significant amounts of analysis. I wanted to analyze the statistics and results from games from the past five years to see if a simple model could outperform the complex analysis being run by the sports books.

Acknowledgments

I would like to thank Professor Gary Dean for advising me throughout my thesis. He has had a major positive impact both on this project, and my entire college career. Without him, this program and my academic success would not be the same.

I would also like to take a moment to thank my parents, sisters, and friends for encouraging me throughout this process.

Table of Contents

Abstract	2
Acknowledgments	3
Process Analysis Statement	5
Introduction	11
Models for Comparison	12
Data	14
Generalized Linear Models	16
Results	19
Conclusions	22
Documentation of Sources	24
Appendix	26

Process Analysis Statement

While I have preferred to watch and follow division one college football my whole life, I choose to stick with the NFL because there is more parity in the professional game. The best team playing the worst team in the NFL will still provide good information to use, while watching the second and third stringers of a top ten college team beat up on a school in the Football Championship Series (FCS) provides almost no predictive value for the rest of the season. Even though a MAC school may average 300 passing yards a game, this does not mean that they will be favored against a Big Ten school that averages 200 total yards a game. These problems were compounded by the fact that in college football the first four games are often used as warm up games for the best teams, providing little data to prepare for much more competitive conference matchups. In addition to all of that, teaching the model to account for massive changes in performance year over year would be very difficult. The NFL also offers four preseason games where some viable data can be collected and used to provide probabilities out of the gate, whereas the NCAA hands you the first week of games with no preparation beyond spring practice with teams where it is not uncommon for half of the starters from the previous season to be replaced. On top of all of those problems, losing one game can derail a college team more completely than an NFL team. We have to look no farther than the Florida State Seminoles last year to see a team lose their first game and go from being title contenders to spiraling to a .500 record.

This is not to say that the NFL does not have similar problems. Franchises can let plenty of key starters leave in the offseason due to trades, retirements, and letting contracts expire. But, teams tend to stay more stable in the NFL than the NCAA, often replacing a few starters

instead of watching their entire set of defensive backs leave for the draft. The largest positive for the NFL is that it is the top tier of football, so these men are the very best from all the college teams that I would have been looking at.

There were two main issues that cropped up: the limited data available and creating models that accurately reflect the data without handcuffing it. The NFL keeps much of the data that it records throughout games to itself, often not even releasing the information to its member teams. This limits how much modeling fans can do as they either must curate the data or pay for the data from a provider. The other issue is the world of 'big data' that we are moving into. A decision must be made by actuaries about the intensity of the modelling they will be doing. This is an argument that I have heard in both of my past internships. Based on the decisions of the companies, I elected to follow their lead, and stick with simpler models that allowed the data to talk.

As my inspiration for the models that I wanted to build, I focused on arguments that I have often heard as to how to create a winning football team. The prevailing themes were: building a dominate defense, a spread offense that used the entire field, a ball control offense that would tire the opponents, spending exorbitant amounts of money for the best players, and using advanced statistics to better determine the most effective players. Once I came to this decision, the problem with poor data availability reared its head. This is often a problem with actuaries as data is not available or it is inaccurate.

This lack of data actually resulted in the destruction of two models: the spread offense model, and the fiscally based model. The first to go was the fiscal model because there was not nearly enough clarity on contracts and annual staff expenditures for any teams. The issue with

the spread offense model was the fact that I wanted to focus on stats that represent a well-run spread attack. There is not quality tracking of the more advanced statistics such as percentage of plays run outside the numbers in the NFL. In order to retrieve this data stream, I would have had to watch every game to record the number of plays where the ball got outside the numbers on the field.

The next hurdle I had to jump was home field advantage. Luckily, every team plays eight home and eight away games in a given regular season, so the per game averages would house the teams average on a neutral field. I still accounted for the home field advantage that exists in the NFL by including it on the back end of my model, rather than the predictive front end. I made this decision for clarity within the model, and to eliminate any bias against bad teams had it been included on a per team basis. The next issue was data cleaning as it had to be put in a format that would work well with the regression software that I decided to use. Luckily, it is a user-friendly software, so most of it was manipulation to create meaningful variables.

EVIIEWS is a very powerful econometric software that could run the regression easily. I realized that I would need to run the regression twice for each model. The first run was simply to determine which of the variables had statistical significance, and the second was to find the correct coefficients for those material variables. Once I started the regressions, I also decided to attempt offsetting the data by one season to see if the statistics from the previous year held any predictive power for the next season. The results made it abundantly clear that the previous year's averages held no predictive power due to the massive amount of change each offseason.

One major issue with utilizing EVIEWS was that I only had the Student Version where no workbooks or data sets could be saved. I had to ensure that I kept all of the data safe in an Excel file, in case I needed to go back and rerun the regression at any time in the future. This did come in handy when it was pointed out that I should switch from a traditional linear regression to a generalized linear model. At that point, I had to go back into EVIEWS and redo all the work that I had previously done.

Once all of the regressions had been run, the number of predictive variables had been pruned considerably. The second model that I had created held just one variable with predictive power, which made me concerned for the efficacy of that model. I still wanted to see it through, without holding much hope, so I decided to go ahead and use it despite the fact that it almost assuredly did not have the ability to compete with my other models, much less the lines that Vegas made available.

The regression for the combination model was the most surprising for me. Once I dove into the regression a little bit, I was able to see why it only had four statistically significant variables. Some variables that had been included in one of the previous three models were excluded because they were correlated with others that were included. Those took some of the explanatory power from the variables who did not have to compete with them in the smaller models. This resulted in the four best variables stepping out and being included in the final combination model, even though seven different variables had proven useful at other times in the process.

Then came the struggle of the unreliable preseason data. Oftentimes, teams run limited playbooks and give their starters limited playing time in the preseason. This worried me as to

how effective the full game averages from the preseason would be. However, those games were the only ones to provide predictive power for the first week of the season, so I decided to move forward with using the full averages from all four games. This turned out to be a great decision as it resulted in competitive models for weeks one through three of the regular season. After that point, the biggest hindrance to my models was the lack of new data being added. While both Nate Silver's 538 model and Vegas were learning from the results, my models were kept blind.

My models were kept blind due to the time constraints on my thesis and my own time. While I wanted to continue to update the data streams, I did not have the time available to dedicate to my thesis given everything else that I needed to do. As the week 1 games were ending, I found myself in the heart of interview season having a minimum of two interviews a week. I believe this was the single biggest challenge that I have faced throughout this project. I have been unable to manage my time to ensure that I could do all of my work, interviews, and executive board responsibilities at the level that I wanted to. I quite simply did not have time to go through the data extraction and cleaning at a level that matched what I had been doing for my model.

Continually updating the data streams is the single biggest way to improve my models for future use. If I try to once again to use the models next season, I will attempt to do three things. The first would obviously be to update the data inputs between every game. This would allow the per game averages to better represent each team as the season went on. The second would be to re run the regression with the results from the 2018 season included. This would expand the years of data that the regression is based on from five to six, hopefully improving

the results. The final change would be to compare the newly regressed models with this year's versions throughout next season. I imagine that the coefficients and results would be similar, but I am curious to see how much of an impact one more year of data makes on the model.

Introduction

The NFL has been around entertaining Americans since the early 1900's and people have been betting on it since its inception. As such, people have always tried to predict who would win, and how likely they were to win that game. Naturally, the game has evolved significantly since then, looking almost nothing like what was originally played. Stretching from uniforms to the forward pass it is almost unrecognizable. However, some people have been able to stay on top of those changes and consistently beat those taking bets. This has grown significantly harder as data was collected, and computers were introduced to model games. Those sports books in Vegas are getting better at making the lines tougher to take advantage of a mistake. To combat this, those who make their money off of betting on sports have also taken to advanced statistical modeling. They are constantly monitoring lines as they change and placing bets on small discrepancies that they find. The goal of this thesis is to find a way to outperform the sports books with simple models.

In order to outperform the big betting books, it is necessary to compare the actual results against both the probabilities that are implied by money lines set by Vegas and the resultant percentages that are produced by the generalized linear model that will be created. In order to do this, the betting lines had to be converted to their implied probability. Since each game involves two teams, the home teams would be tracked with the implication being that tracking both would not add any more information as the probability of the away team winning is one minus the probability of the home team winning. Thus, if the home team wins, all models being compared would receive the probability they predicted that the home team would win, and if the away team won, they would gain the probability they predicted that the away team

would win. In ties, each model would receive half of the probability that they predicted the home team would win.

In order to compare the effectiveness of simplicity, it was also decided to use the same method of comparison against another statistical model created by Nate Silver. He uses a more holistic statistical methodology to create a predictive model for the NFL (Silver, 2018). Instead of looking at any specific statistics, he takes the full results of the game and changes the teams overall rating to better predict the future outcomes. This would be used to test if specific statistics could be used to more accurately predict winners than overall performance.

Models for Comparison

As stated above, Nate Silver's ELO model and Vegas' betting lines would be used to compare against my individually based statistical model. While Silver and his staff at 538 publish their methodology, this is not done by any major Vegas sportsbook. This makes sense when one considers the fact that the sportsbooks make their money off of people betting on both sides of their lines whereas 538 makes money through people interested in statistics who visit their site and read their articles. The difference in income streams justifies the distribution of information from each source.

It is important to mention that Vegas sportsbooks set lines to draw even action on both sides, as stated above. This means that they are focused on where the largest bettors believe the line should be rather than what is actually most likely to occur. Oftentimes, this results in a line that is reasonable to most people, and is close to the theoretical odds that would be produced without bias towards or against one team. This all means that the comparison to

Vegas will actually be a way to judge the betting, and often modeling, public on how they believe games will turn out. Thus, the way to turn a profit off of Vegas is to find the flaws in the betting logic of the public.

Diving more into Nate Silver's ELO model, one sees that he does pull results from the previous season (Silver, 2018). However, he recognizes that teams do not return at the same level that they left the previous season at, and regresses every team towards the mean in order to reflect the gains and losses through free agency, the draft, retirements, trades, and coaching changes. Once the preseason begins, the ELO rating changes based primarily on the margin of victory and a k factor created to limit how much the rating fluctuates based on results. The k factor was determined by looking at historical data in a manner not disclosed to the public. The margin of victory gets treated to a natural log function, which decreases the weight for each extra point the team won a game (Silver, 2018). So, the first couple points a team wins by carry the most weight, while winning by 41 or 42 does not change a team's ELO by very much. This methodology carries the advantages of not needing to adjust inputs for massive changes in the game. The introduction of new rules, offenses, and defenses will not hinder the model from reacting to the results that come from them. In those cases, it would often need a few weeks to account for the advantages they may offer certain teams, but it would get there relatively quickly without requiring the creation of a new model.

As I briefly discussed earlier, I decided to compare my models predictions with that of Vegas odds makers and 538 sports website. While they occasionally have bad weeks it is well known that Vegas always ends up making money. So, if my model could beat the odds that Vegas puts out, then I would consider it a highly successful and effective model. The way to

compare my percentages to Vegas's betting lines required finding the implied probabilities in each spread. Luckily, some bettors have already created a convenient chart that allowed me to convert the line to an implied probability that either team would win ("What is the Percentage...", 2018). I also decided to also include Nate Silver and his model on his website, 538.com. I wanted to compare my model against theirs because it has done an excellent job outperforming their reader base. People can input their own probabilities that each team will win, and the results are posted for all to see. 538's ELO model falls in the 93rd percentile, beating the vast majority of people who try to pick winners (Boice, Bycoffe, & Wezerek, 2018). The other reason I wanted to include this model was to see if my models simplicity could beat the more complicated model that is based on more advanced statistics.

Data

Data was accumulated dating back five years. The results from every regular season game from 2013-2017 were included as information for the basis of the regression of the models that were created. The reason for collecting only five years of data was due to the large number of rule changes, along with the increased acceptance of spread concepts in professional football. Offenses lining up under center and running the ball for three yards a play no longer provide useful information in 2018 as shotgun sets have grown massively in popularity. As discussed in the process analysis statement, three models were created along with a fourth that was simply a combination of the other three. For each of them, data was collected from ESPN.com as it had been curated back well before the 2013 line.

The statistics for model one include: opponent's yards per game, opponent's points per game, turnovers per game, and penalties per game.

The statistics for model two include: plays run, yards per play, quarterback rating, and fourth down conversion percentage.

The statistics for model three include: points per play, success rate, adjusted points per play, average punt length, and turnover margin.

The statistics for the final model were: opponent's yards per game, opponent's points per game, turnovers per game, penalties per game, plays run, yards per play, quarterback rating, fourth down conversion percentage, points per play, and average punt length. Success rate, adjusted points per play, and turnover margin were all eliminated for directly correlating with another variable.

Points per play is the number of points earned for every play an offense runs. Success rate is the percent of plays where the offense was able to pick up at least half of the necessary yards for a first down, or converting on third down. Adjusted points per play is the number of points earned for every play adjusted based on the yard line the play started on. This is due to the fact that plays inside the red zone are worth more than plays on a team's own 20-yard line.

Naturally, not all of these statistics were directly available on ESPN.com. For those that involved the combination of multiple statistics, the equations are available in the appendix. The other main problem within the data was struggling to pin down how to utilize statistics such as quarterback rating. This stream in particular was a problem because some teams use multiple quarterbacks in game due to injury or performance. In order to properly represent how

quarterbacks played, I decided to weight each quarterback's average rating by the number of quarters they played.

Of course, in any sporting event a team can win, lose, or tie. The number that I was interested in predicting was wins, so wins from the previous five seasons were what was recorded and used as the output. This meant that every game was either a win or not a win, which was recorded as a one or a zero respectively. Since this does not show the actual probability that a team wins, I knew from collegiate classes that it ought to be viewed as a binomial distribution with 16 occurrences.

Generalized Linear Models

The method that was utilized to create the three models was a generalized linear model. Many people first think of a traditional linear model, which looks like:

$$Y = \beta_0 + \beta_1 * X_1 + \dots + \beta_n * X_n$$

However, this model has issues when attempting to predict the winner of a game. A traditional linear model has a range from negative infinity to positive infinity, while a probability has a range from zero to one. However, this linear model can be generalized, and the dependent variable, Y , be plugged into another function. In order to adjust the range of the output, I decided to use a logit link function. The first step of this is to create the function to replace Y above. For this, one uses $g(y)=\ln(y)$, which can be plugged into the equation. The new model looks like:

$$\ln(Y) = \beta_0 + \beta_1 * X_1 + \dots + \beta_n * X_n$$

Now, this can also be represented as $Y = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}$. This takes care of the lower bound of infinity by bringing it to zero as the exponential function can only be positive. In order to ensure that the upper bound is one, the logit function comes in. To do this, we divide the above function by one plus itself, which will ensure it never becomes greater than one. So, the final equation used was:

$$Y = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}}$$

Now, the model could be fitted to the data while always producing a result between zero and one. This process was made much easier since the software that I decided to use was already built to handle generalized linear models through a logit link function to a 16 count binomial data set. The EVIEWS software was extremely powerful and user friendly to run the data analysis, requiring almost no cleaning or adjusting of the data to run properly.

Shown in the appendix is a sample of the printout that EVIEWS provides after running regression. In order to keep the models relevant and as accurate as possible, any coefficient with a p-value greater than .05 would be removed with the remaining coefficients run again to find the accurate value for each variable. When a regression is run, the coefficient is essentially never zero, however, some coefficients ought to be zero because the variable is not predictive. The p-value is the corresponding probability that a coefficient is supposed to be zero. So, if a p-value is less than .05, then there is less than a 5% chance that the variable actually does not have predictive power for the model. I chose to eliminate variables along these lines to keep the data clean and results as potent as possible.

Due to this elimination, the final models that the new data was plugged into used only the following factors:

- The statistics for model one include: opponent's points per game, and turnovers per game.
- The statistics for model two include: quarterback rating.
- The statistics for model three include: points per play, success rate, adjusted points per play, average punt length, and turnover margin.
- The statistics for the final model were: opponent's points per game, turnovers per game, quarterback rating, and points per play.

While it was not ideal to see only two variables affect the first and one affect the second, it is better than including variables that do not hold predictive power. This did lead me to believe that the first two would be poorer predictors of the games than the third or combination models. The fact that the third model did not lose a single variable gave me hope that it could potentially be the best model, and possibly compete with Vegas and Nate Silver's model.

Since I regressed the model based on a binomial distribution with 16 occurrences, it would actually produce an expected number of wins on the season, rather than a percentage chance to win a single game. Those would be against an average team, so by dividing by 16, I could extract the percentage chance of beating a completely average team, which I called the 'expected wins per game'. Since, in any given week, each team does not play another who is considered completely average, I had to combine the results for both teams playing each other. The challenge here was how to accurately combine them to find the win probability for each team on a neutral field. My solution ended up being pretty simple. I took the home teams expected wins per game and divided it by the sum of both teams expected wins per game. This

was ensured to provide a probability that the home team won, and adjusted it based on the skill of the other team. If their opponent was significantly better, then it would create a larger denominator resulting in a smaller probability. All that was left was to include the home field advantage factor discussed earlier.

Results

The largest fear I had early in the process was the fact that no data from last season could be used to predict how well a team would do in the next season. This made the four preseason games for each team the only input for week one. While this may not sound so bad initially, oftentimes starters only play for a quarter of each preseason game, with playbooks being limited so as not to give away too much to a team's week one opponents. As discussed in the process analysis statement, these ended up being the only inputs that the model had available for every week. I believe this greatly affected all of the model's abilities to compete in the later weeks.

When this project was undertaken, it was decided to track only the first eight weeks of the regular season so as to provide time to write the thesis paper. As described earlier, each model would get credit based on the probability they assigned to the winning team. With this system, each model can receive a score from zero to 16 every week. This means that if a model gave every team a 50% chance of winning then it would earn 8 points every week without byes. Since 14 teams had byes through week 8, a model assigning 50% to every team would result in a total score of 57.0 (NFL Schedule, 2018).

Even the worst model I had created outperformed that as model two earned an aggregate score of 63.30. Despite outperforming model two in every week along with being my most accurate model, model three only produced an aggregate score of 64.84. The total results along with the week by week results can be found in the appendix. The highest score each week has a light green highlight while the lowest score has a light orange highlight. This relatively tight range of scores falls significantly below that of both Nate Silver's model and the lines set by Vegas sportsbooks, scoring 67.60 and 70.23 respectively.

The largest issue for my models in keeping up with the others was the lack of new inputs as each week of games were resolved. Instead of adding new data and learning from each passing game, my models were stuck in the preseason, blind to the changes that were occurring as teams began to build chemistry or fall apart. The range for all six models being compared each week can be found in the appendix, and it shows a steadily increasing gap between the best and worst models. While there were blips in weeks one and four, there appears to be a correlation between how late in the season the games were and how closely the models performed. It is also important to point out that in the first four weeks, two were marked by Vegas sportsbooks or Nate Silver having the best week, while in the last four weeks, all of them were led by one of those two.

There was a similar trend for the worst performer of each week. Three out of the first four had Vegas or 538 performing the worst while none of the last four had them in the bottom half of performers. One can see in the appendix that all four of my models remained roughly as good at predicting the winners throughout all eight weeks. While there may initially appear to be a decrease in performance, byes started in week four decreasing the maximum number of

points from 16 to 15. Then, in weeks seven and eight, the maximum number was decreased again from 15 to 14 as the number of teams with byes were doubled from two to four. This makes Vegas's week eight results that much more impressive as they scored 9.155 out of 14 total possible points. Three key games for them were Pittsburgh, Kansas City, and the LA Rams netting them a total of 2.438 points by winning.

As expected when all but one of the variables turned out to lack predictive power, model two performed the worst taking home three of the eight worst weeks, cracking the top half of performers only once, placing third. While it did beat calling every game a coin toss, it did not perform well, being the only model below 64 total points out of the 114 available. The points earned can also tell another story. Based on the fact that Vegas earned 70.2275 points, they earned an average of .6160 points per game. This would mean that they gave the average winner a probability of 61.6% chance of winning. My models tended not to be as bold, which resulted in lower overall scores.

This actually gives me a little bit of hope as this has been a year of close matchups. There have been a fair amount of upsets, a lot of close matchups, and only a few teams who have stuck out as better than most of their competition (Paine, 2018). These results would tend to imply that a model that is less certain has some validity to it. However, at the outset I decided that I would compare my model results against the lines placed by Vegas and picking the winners with more certainty than Nate Silver's 538 model. By using this method of comparison, there is no reward for focusing on parity and punishing for overconfidence. This is directly opposed to the scoring system set out by the 538 website, which punishes

overconfidence on wrong predictions (Boice, 2018). However, it was chosen to compare against Vegas spreads and the 538 predictions without a graduated scale.

Conclusion

As you can see in the results above, the models performed well early in the season while the preseason data was still relevant, but quickly lost steam as both Vegas and Nate Silver continued to update their models with new data. Due to these revelations, it cannot be concluded whether these generalized linear models can be competitive with Vegas or Nate Silver over the course of a whole season. The first few weeks offer some hope, but both competitors adjust their expectations heavily based on results from the season so far. The improvement in performance for Vegas and Nate Silver is due to them both accounting for wins that teams have earned so far. It is interesting to note that their ability to correctly predict games improved, while my models actually did not deteriorate.

I would have expected my models to actually worsen in their predictive capabilities as the data grew more and more outdated. However, they continued to provide roughly the same results as the weeks passed. Shown in the Appendix, both Vegas and Nate Silver's 538 model were able to improve as all of mine tended to stay reasonably consistent. I think part of this could be due to the reasonably close percentages that my models offered to all teams. Early on in the season, this proved to be a boon as some teams who were worse last year and had improved quite a bit. My models benefitted from the fact that some worse teams from last season performed better than Vegas and Nate Silver's model expected early in the year. As the

year drug on my models were hindered by their inability to choose a clear favorite in some of the more lopsided matchups.

Both Vegas and 538 held more confidence in the teams they predicted to win than any of my models. I believe that this could be a good thing as there are a lot of competitive games, and many of them could have gone either way. That said, my models did not perform well given the metrics that I had laid out in advance. If my models had continued to receive their inputs, then they might have become more confident. However, I am not convinced that they would have become more confident in their predictions. The results were not enough to draw conclusions that my models were better than Vegas or Nate Silver's model. The first three weeks provided optimism while the last five showed the weaknesses in a more moderate modeling system.

References

- Boice, J., Bycoffe, A., & Wezerek, G. (2018, November 06). Can You Beat FiveThirtyEight's NFL Forecasts? Retrieved November 7, 2018, from <https://projects.fivethirtyeight.com/2018-nfl-forecasting-game/leaderboard/>
- Connelly, B. (2014, January 24). College football's 5 most important stats. Retrieved March 23, 2018, from <https://www.footballstudyhall.com/2014/1/24/5337968/college-football-five-factors>
- NFL point spreads: Weekly NFL point spread - Free Vegas NFL spreads. Pro football betting spreads. (n.d.). Retrieved from http://www.footballlocks.com/nfl_point_spreads.shtml
- NFL Schedule - 2018. (n.d.). Retrieved from <http://www.espn.com/nfl/schedulegrid>
- NFL Team Total Offense Statistics. (n.d.). Retrieved from http://www.espn.com/nfl/statistics/team/_/stat/total
- NFL Total QBR. (n.d.). Retrieved from <http://www.espn.com/nfl/qbr>
- Paine, N. (2017, May 12). A Home Playoff Game Is A Big Advantage - Unless You Play Hockey. Retrieved June 18, 2018, from <https://fivethirtyeight.com/features/a-home-playoff-game-is-a-big-advantage-unless-you-play-hockey/>
- Paine, N. (2018, October 10). NFL Parity Is Creating Excitement - And, Um, Ties. Retrieved October 24, 2018, from <https://fivethirtyeight.com/features/nfl-parity-is-creating-excitement-and-um-ties/>
- Silver, N. (2018, September 05). How Our NFL Predictions Work. Retrieved October 21, 2018, from <https://fivethirtyeight.com/methodology/how-our-nfl-predictions-work/>

What is the percentage chance a team wins the game based on the closing point spread? (2018, May 29). Retrieved June 18, 2018, from <http://www.bettingtalk.com/win-probability-percentage-point-spread-nfl-nba/>

Appendix

Statistic Creation:

Plays Run	Rushing Attempts + Passing Attempts
Yards Per Play	Total Yards / Plays Run
Points Per Play	Total Points / Plays Run
Adjusted Points Per Play	Points Per Play / Avg. Starting Field Position

Sample EVIEWS Printout:

Dependent Variable: WINS

Method: Generalized Linear Model (Newton-Raphson / Marquardt steps)

Date: 09/05/18 Time: 13:22

Sample: 1 160

Included observations: 160

Family: Binomial Count (n = 16)

Link: Logit

Dispersion fixed at 1

Summary statistics are for the binomial proportions and implicit variance weights used in estimation

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
OPP_YPP	0.003220	0.002186	1.473262	0.1407
OPP_PPG	-0.114898	0.020823	-5.517877	0.0000
TO_PG	0.408280	0.096770	4.219052	0.0000
PENALTIES	-0.025666	0.044909	-0.571518	0.5676
C	1.641870	0.632415	2.596192	0.0094
Mean dependent var	0.003115	S.D. dependent var	0.048563	
Sum squared resid	3.730208	Log likelihood	-378.1027	
Akaike info criterion	4.788784	Schwarz criterion	4.884883	
Hannan-Quinn criter.	4.827807	Deviance	265.6776	
Deviance statistic	1.714049	Restr. deviance	414.9714	
LR statistic	149.2938	Prob(LR statistic)	0.000000	
Pearson SSR	255.7830	Pearson statistic	1.650213	
Dispersion	1.000000			

Week by Week Results:

	Vegas	538	Model 1	Model 2	Model 3	Combo
Total	70.2275	67.595	64.30474	63.30165	64.84378	64.63282
Week 1	8.9565	8.615	8.056446	7.911642	8.141247	8.023966
Week 2	8.777	8.38	8.381824	8.762215	8.840497	8.756678
Week 3	8.024	8.26	8.775776	8.181033	8.422354	8.64903
Week 4	9.242	8.45	7.959898	7.809188	8.196244	7.950976
Week 5	8.866	8.42	8.534022	7.932268	8.244274	8.336601
Week 6	8.911	8.97	7.766277	7.902043	7.959615	7.746234
Week 7	8.296	7.68	7.037649	7.207876	7.337394	7.178385
Week 8	9.155	8.82	7.792851	7.595387	7.702153	7.990953

Range of Results:

	RANGE	PERCENT CHANGE
WEEK 1	1.044858	
WEEK 2	0.460497	-55.927%
WEEK 3	0.751776	63.253%
WEEK 4	1.432812	90.590%
WEEK 5	0.933732	-34.832%
WEEK 6	1.223766	31.062%
WEEK 7	1.258351	2.826%
WEEK 8	1.559613	23.941%

Average Percent Chance Given to Actual Winners:

	VEGAS	538	MODEL 1	MODEL 2	MODEL 3	COMBO
WEEK 1	56.0%	53.8%	50.4%	49.4%	50.9%	50.1%
WEEK 2	54.9%	52.4%	52.4%	54.8%	55.3%	54.7%
WEEK 3	50.2%	51.6%	54.8%	51.1%	52.6%	54.1%
WEEK 4	61.6%	56.3%	53.1%	52.1%	54.6%	53.0%
WEEK 5	59.1%	56.1%	56.9%	52.9%	55.0%	55.6%
WEEK 6	59.4%	59.8%	51.8%	52.7%	53.1%	51.6%
WEEK 7	59.3%	54.9%	50.3%	51.5%	52.4%	51.3%
WEEK 8	65.4%	63.0%	55.7%	54.3%	55.0%	57.1%